Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 429 (2008) 1026–1037
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Continuity properties of the {1}-inverse and perturbation bounds for the Drazin inverse ☆

Xin-guo Liu [a], Wei-guo Wang [a], Yi-min Wei [b,c,*]

[a] *Department of Mathematics, Ocean University of China, Qingdao, Shandong 266071, PR China*
[b] *School of Mathematical Sciences, Fudan University, Shanghai 200433, PR China*
[c] *Key Laboratory of Mathematics for Nonlinear Sciences (Fudan University), Ministry of Education, Shanghai 200433, PR China*

Received 31 October 2006; accepted 22 March 2007
Available online 11 April 2007
Submitted by H. Bart

Abstract

There are known results about the continuity of the important pseudoinverses. We extend some well known results to the {1}-inverse. As for the Drazin inverse, which is not a {1}-inverse, we present new perturbation bounds.

AMS classification: 15A09; 65F20

Keywords: {1}-Inverse; Oblique projection; Continuity; Drazin inverse

1. Introduction

Much attention has been paid to the continuity of the important pseudoinverses. For instance, Wedin [28] and Stewart [25] studied the Moore–Penrose inverse; Wang [27] investigated the weighted Moore–Penrose inverse; Campbell and Meyer [2–4] established a necessary and sufficient condition for the continuity of the Drazin inverse.

☆ The 1st and 2nd authors are supported by NSF of Shandong province (Y2000A04), the third author is supported by the National Natural Science Foundation of China under Grant 10471027 and Shanghai Education Committee.
* Corresponding author. Address: School of Mathematical Sciences, Fudan University, Shanghai 200433, PR China.
E-mail address: ymwei@fudan.edu.cn (Y.-m. Wei).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2007.03.032

Let $A \in \mathbf{C}^{m\times n}$ be given. The matrix $X \in \mathbf{C}^{n\times m}$ satisfying $AXA = A$ is called a $\{1\}$-inverse of matrix A and has many applications [1]. Some researchers have considered the $\{1\}$-inverse (i.e., [1]). The symbol $A\{1\}$ denotes the set of $\{1\}$-inverse of A. $\mathrm{rank}(A)$ denotes the rank of A. $\mathscr{R}(A)$ and $\mathscr{N}(A)$ represent the range space and null space of A, respectively. $\|A\|_2$ is the spectral norm of A.

There are two issues considered in this paper. Firstly, we know that the Moore–Penrose inverse, $\{1, 3\}$-inverse, $\{1, 4\}$-inverse, weighted Moore–Penrose inverse and the group inverse are all $\{1\}$-inverse; we consider the continuity of the $\{1\}$-inverse and associated oblique projections in Section 2. Secondly, we also notice that the continuity of the Drazin inverse not a $\{1\}$-inverse has been investigated by Campbell and Meyer [3]. Recently, several other authors have developed perturbation bounds of the Drazin inverse [5,6,7,8,9,10,11,14,15,16,17,19,20,22,29,30,31,32,33,34,35].

However, we believe that further exploration about the perturbation bounds are important. In Section 3, some examples are presented to illustrate the known results, new perturbation bounds for the Drazin inverse A^{D}, and the associated oblique projection.

2. Continuity of the {1}-inverse and associated oblique projections

In this section, we will investigate the continuity of the $\{1\}$-inverse and its oblique projections. First, applying an elegant theorem of Wedin [28], we can prove the following result.

Theorem 2.1. *Let $A, B = A + E \in \mathbf{C}^{m\times n}$, $A^- \in A\{1\}$. If $\mathrm{rank}(A) < \mathrm{rank}(B)$, then*

1. *For any $B^- \in B\{1\}$, we have*
$$\|B^-\|_2 \geqslant \frac{1}{\|E\|_2}; \tag{2.1}$$
2. *If $\mathscr{R}(A) = \mathscr{R}((A^-)^H)$, then*
$$\|B^- - A^-\|_2 \geqslant \frac{1}{\|E\|_2}. \tag{2.2}$$

Proof. Since $\mathrm{rank}(A) < \mathrm{rank}(B)$, there is a $y \in \mathscr{R}(B)$ such that
$$\|y\|_2 = 1, \quad y^H A = 0.$$
Noting that BB^- is the oblique projector along $\mathscr{N}(B)$ onto $\mathscr{R}(B)$, we get
$$1 = yy^H = y^H BB^- y = y^H EB^- y \leqslant \|E\|_2 \|B^-\|_2.$$
Then the inequality (2.1) follows.

Moreover, if $\mathscr{R}(A) = \mathscr{R}((A^-)^H)$, then $y^H (A^-)^H = 0$,
$$1 = y^H BB^- y = y^H E(B^- - A^-)y \leqslant \|E\|_2 \|B^- - A^-\|_2. \qquad \square$$

The following result gives a new expression of A^- in terms of the Moore–Penrose inverse and the full rank decomposition.

Lemma 2.2. *Let $A \in \mathbf{C}^{m\times n}$, $\mathrm{rank}(A) = r > 0$, and $A = BC$ be the full rank decomposition. Then*
$$A\{1\} = \{C^\dagger B^\dagger + Z - C^\dagger CZBB^\dagger | Z \in \mathbf{C}^{n\times m}\}, \tag{2.3}$$
where $B^\dagger$ and $C^\dagger$ are the Moore–Penrose inverse [1,2,27] of B and C, respectively.

Proof. By the decomposition $A = BC$, $B \in \mathbf{C}_r^{m\times r}$, $C \in \mathbf{C}_r^{r\times n}$, we have

$$\begin{aligned} AXA = A &\Leftrightarrow BCXBC = BC \\ &\Leftrightarrow CXB = I_r. \end{aligned}$$

Following a result due to Sun [24, Lemma 1.3], we obtain

$$A\{1\} = \{X = C^\dagger B^\dagger + Z - C^\dagger CZBB^\dagger | Z \in \mathbf{C}^{n\times m}\}. \qquad \square$$

Now we consider the continuity of the $\{1\}$-inverse as follows.

Theorem 2.3. *Let* $A, B \in \mathbf{C}^{m\times n}$, $B = A + E$. *For each* $A^- \in A\{1\}$, *if* B *approaches* A, $\operatorname{rank}(A) = \operatorname{rank}(B)$, *then there exists* $B_A \in B\{1\}$, *such that*

$$\lim_{B\to A} B_A = A^-. \tag{2.4}$$

Proof. Suppose the singular value decomposition (SVD) [13] of A is $A = U\begin{pmatrix}\Sigma_1 & 0\\ 0 & 0\end{pmatrix}V^H$, where U and V are unitary matrices, $\Sigma_1 = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r) > 0$. It is easy to check that

$$A^- = V\begin{pmatrix}\Sigma_1 & 0\\ 0 & 0\end{pmatrix}^- U^H.$$

Since

$$\begin{pmatrix}\Sigma_1 & 0\\ 0 & 0\end{pmatrix} = \begin{pmatrix}\Sigma_1\\ 0\end{pmatrix}\begin{pmatrix}I_r & 0\end{pmatrix},$$

it follows from Lemma 2.2 that

$$\begin{pmatrix}\Sigma_1 & 0\\ 0 & 0\end{pmatrix}^- = \begin{pmatrix}I_r & 0\end{pmatrix}^\dagger \begin{pmatrix}\Sigma_1\\ 0\end{pmatrix}^\dagger + Z_1 - P_1 Z_1 P_2,$$

where

$$P_1 = \begin{pmatrix}I_r & 0\end{pmatrix}^\dagger \begin{pmatrix}I_r & 0\end{pmatrix}, \quad P_2 = \begin{pmatrix}\Sigma_1\\ 0\end{pmatrix}\begin{pmatrix}\Sigma_1\\ 0\end{pmatrix}^\dagger.$$

Suppose that A^- is a $\{1\}$-inverse of A. Then there exists a $Z_1 \in \mathbf{C}^{n\times m}$ such that

$$A^- = V\left\{\begin{pmatrix}I_r & 0\end{pmatrix}^\dagger \begin{pmatrix}\Sigma_1\\ 0\end{pmatrix}^\dagger + Z_1 - P_1 Z_1 P_2\right\}U^H. \tag{2.5}$$

The expression $B = A + E = U\left\{\begin{pmatrix}\Sigma_1 & 0\\ 0 & 0\end{pmatrix} + E_0\right\}V^H$ implies

$$B^- = V\left\{\begin{pmatrix}\Sigma_1 & 0\\ 0 & 0\end{pmatrix} + E_0\right\}^- U^H.$$

Let E_0 be partitioned as

$$E_0 = \begin{pmatrix}E_{11} & E_{12}\\ E_{21} & E_{22}\end{pmatrix}.$$

If $\operatorname{rank}(A) = \operatorname{rank}(B)$ and $\|E\|_2$ is sufficiently small, then we have

- $B_{11} \equiv \Sigma_1 + E_{11}$ is nonsingular.
- $\operatorname{rank}\begin{pmatrix} B_{11} & E_{12} \\ E_{21} & E_{22} \end{pmatrix} = \operatorname{rank}(B_{11})$.
- $\begin{pmatrix} B_{11} & E_{12} \\ E_{21} & E_{22} \end{pmatrix} = \begin{pmatrix} B_{11} \\ E_{21} \end{pmatrix} \begin{pmatrix} I_r & F_{12} \end{pmatrix}$, where $F_{12} = B_{11}^{-1} E_{12}$.

It follows from Lemma 2.2 that

$$\begin{pmatrix} B_{11} & E_{12} \\ E_{21} & E_{22} \end{pmatrix}^{-} = \begin{pmatrix} I_r & F_{12} \end{pmatrix}^{\dagger} \begin{pmatrix} B_{11} \\ E_{21} \end{pmatrix}^{\dagger} + K - \widetilde{P}_1 K \widetilde{P}_2,$$

where $\widetilde{P}_1 = (I_r \quad F_{12})^{\dagger}(I_r \quad F_{12})$, $\widetilde{P}_2 = \begin{pmatrix} B_{11} \\ E_{21} \end{pmatrix}\begin{pmatrix} B_{11} \\ E_{21} \end{pmatrix}^{\dagger}$, and K is arbitrary.

Let $K = Z_1$, we obtain

$$B_A = V \left\{ \begin{pmatrix} I_r & F_{12} \end{pmatrix}^{\dagger} \begin{pmatrix} B_{11} \\ E_{21} \end{pmatrix}^{\dagger} + Z_1 - \widetilde{P}_1 Z_1 \widetilde{P}_2 \right\} U^H. \tag{2.6}$$

Notice that $E \to 0$ which implies

$$F_{12} \to 0, \quad E_{12} \to 0, \quad B_{11} \to \Sigma_1, \quad B_{11}^{-1} \to \Sigma_1^{-1}.$$

Thus

$$\widetilde{P}_1 \to P_1, \quad \widetilde{P}_2 \to P_2.$$

It follows from (2.5) and (2.6) that

$$\lim_{B \to A} B_A = A^{-}. \qquad \square$$

Remark 2.4. The upper bound of $\|B_A - A^{-}\|_2$ can be obtained from (2.5) and (2.6).

Remark 2.5. From (2.5) and (2.6), we can see that B_A is the same category of $\{1\}$-inverse as A^{-}.

In the remainder of section, we consider the continuity of the oblique projector AA^{-}.

Theorem 2.6. *Let* $A, B \in \mathbf{C}^{m \times n}$. *If* $\operatorname{rank}(B) > \operatorname{rank}(A)$, *for any* $A^{-} \in A\{1\}$, $B^{-} \in B\{1\}$, *then*

$$\|BB^{-} - AA^{-}\|_2 \geqslant 1. \tag{2.7}$$

Proof. Let $M \in \mathbf{C}^{m \times n}$. Suppose the SVD of M is

$$M = U \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V^H,$$

where U and V are unitary matrices, $\Sigma_1 = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r) > 0$. From $MM^{-}M = M$, we have

$$M^{-} = V \begin{pmatrix} \Sigma_1^{-1} & M_1 \\ M_2 & M_3 \end{pmatrix} U^H.$$

Thus, $MM^- = U\begin{pmatrix} I_r & K \\ 0 & 0\end{pmatrix}U^H$, for any $K \in \mathbf{C}^{r\times(m-r)}$.

Next, using the SVD of K, it is easy to obtain: if the singular values of K are $\sigma_1, \sigma_2, \ldots, \sigma_r$, then the singular values of (I_r, K) are $\sqrt{1+\sigma_j^2}$, $j = 1, 2, \ldots, r$.

Finally, since $\mathrm{rank}(B) > \mathrm{rank}(A)$, we have

$$AA^- = U_1\begin{pmatrix} I_r & K_1 \\ 0 & 0\end{pmatrix}U_1^H, \quad BB^- = U_2\begin{pmatrix} I_s & K_2 \\ 0 & 0\end{pmatrix}U_2^H,$$

and

$$s = \mathrm{rank}(B) > r = \mathrm{rank}(A).$$

Based on the result of the perturbation of SVD [23, p. 199, Theorem 3.11], we obtain

$$\|BB^- - AA^-\|_2 \geqslant 1. \qquad \square$$

Theorem 2.7. *Let* $A, B \in \mathbf{C}^{m\times n}$, $B = A + E$. *For each* $A^- \in A\{1\}$, *if* B *is close to* A *and* $\mathrm{rank}(A) = \mathrm{rank}(B)$, *then there exists* $B_A \in B\{1\}$, *such that*

$$\lim_{B\to A} BB_A = AA^-. \tag{2.8}$$

Proof. With the same notation as in Theorem 2.3, it is not difficult to see that

$$AA^- = U\begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0\end{pmatrix}\left\{\begin{pmatrix} I_r & 0\end{pmatrix}^\dagger \begin{pmatrix} \Sigma_1 \\ 0\end{pmatrix}^\dagger + Z_1 - P_1 Z_1 P_2\right\}U^H. \tag{2.9}$$

Choosing $B_A \in B\{1\}$, such that

$$BB_A = U\begin{pmatrix} \Sigma_1 + E_{11} & E_{12} \\ E_{21} & E_{22}\end{pmatrix}\left\{\begin{pmatrix} I_r & F_{12}\end{pmatrix}^\dagger \begin{pmatrix} B_{11} \\ E_{21}\end{pmatrix}^\dagger + Z_1 - \widetilde{P}_1 Z_1 \widetilde{P}_2\right\}U^H. \tag{2.10}$$

Note that $B \to A$ implies

$$E_{ij} \to 0 \ \ (1 \leqslant i, j \leqslant 2), \quad B_{11} \to \Sigma_1, \quad F_{12} \to 0, \quad \widetilde{P}_1 \to P_1, \quad \widetilde{P}_2 \to P_2.$$

Consequently, we obtain

$$\lim_{B\to A} BB_A = AA^-. \qquad \square$$

In [3,27], it was pointed out that the increasing rank perturbation of A^k disturbed the continuity of Drazin inverse. We will consider the perturbation bounds of the Drazin inverse in next section.

3. Perturbation bounds for the Drazin inverse and associated oblique projection

As the Drazin inverse is a {2}-inverse, not a {1}-inverse, we consider the perturbation bounds of the Drazin inverse in this section. First of all, we recall the definition of the Drazin inverse (see [1,2] for detail).

Drazin inverse and group inverse are very useful because of variousapplications in singular differential and difference equations, Markov chains, iterative methods and numerical analysis were found in the literature [2,12,19,36], etc.

Definition 3.1 [1,2]. Let $A \in \mathbf{C}^{n\times n}$ with $\mathrm{Ind}(A) = k$. If $X \in \mathbf{C}^{n\times n}$ satisfies

$$A^k X A = A^k, \quad XAX = X, \quad AX = XA, \tag{3.1}$$

then X is called the Drazin inverse of A, and denoted by A^{D}.

In particular, when $\mathrm{Ind}(A) = 1$, the matrix X satisfying (3.1) is called the group inverse, and denoted by $X = A^{\sharp}$.

A necessary and sufficient condition for the continuity of Drazin inverse was posed by Campbell and Meyer in 1975 [3]. They also indicated two difficulties in establishing norm estimates for the Drazin inverse. Making use of the relation with the Moore–Penrose inverse, Rong [22] first gave the first order upper bound for Drazin inverse and its condition number.

Much attention has been paid to the perturbation analysis of the Drazin inverse, see [6,7,8, 9,10,15,16,17,19,22,29,30,31,33,32,35]. However, we believe that there are several aspects still open. First, we present some examples.

Example 3.1. Let $A = \begin{pmatrix} J & 0 \\ 0 & N \end{pmatrix}$, $N = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$, $E = \begin{pmatrix} 0 & 0 \\ 0 & aN \end{pmatrix}$, $B = A + E$, where J is nonsingular.

It is easy to see that $B^{\mathrm{D}} = A^{\mathrm{D}}$, and $\|B - A\|_2 = |a|$.

This example shows that the perturbation bounds based on $\|B - A\|_2$ are sometimes overestimates of the perturbation. In addition, results in [29] cannot be applied to this example although we can take $|a|$ small enough. If $|a|$ is not small enough, then the results obtained by Rong [22], Wei and Wu [31] this example will give overestimates.

Example 3.2. Let A be the same as Example 3.1, $N \in \mathbf{C}^{k\times k}$, $E = \begin{pmatrix} 0 & E_0 \\ 0 & 0 \end{pmatrix}$, $B = A + E$. It is easy to see that

$$E_k \equiv B^k - A^k = \begin{pmatrix} 0 & K \\ 0 & 0 \end{pmatrix},$$

where

$$K = \sum_{i=1}^{k-1} J^{k-i} E_0 N^i, \quad AA^{\mathrm{D}} = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad I - AA^{\mathrm{D}} = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}.$$

By direct computation, we have

$$E_k\left(I - AA^{\mathrm{D}}\right) = \begin{bmatrix} 0 & K \\ 0 & 0 \end{bmatrix}, \quad \left(I - AA^{\mathrm{D}}\right) E_k = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This example illustrates that $\|(I - AA^{\mathrm{D}})E_k\|_2$ can be much smaller than $\|E_k\|_2$.

Example 3.3. Let A be the same as Example 3.1, but $E = \begin{pmatrix} 0 & 0 \\ E_0 & 0 \end{pmatrix}$. By some algebra, we obtain $\|E_k(I - AA^{\mathrm{D}})\|_2$ is much smaller than $\|E_k\|_2$.

These two examples show that the quantities $\frac{\|(I-AA^{\mathrm{D}})E_k\|_2}{\|E_k\|_2}$ and $\frac{\|E_k(I-AA^{\mathrm{D}})\|_2}{\|E_k\|_2}$ can be very small.

Example 3.4. Let A be the same as Example 3.1, $E = \begin{pmatrix} 0 & E_0 \\ 0 & 0 \end{pmatrix}$, $B = A + E$. It is easy to obtain that $A^{\mathrm{D}} = \begin{pmatrix} J^{-1} & 0 \\ 0 & 0 \end{pmatrix}$. On the other hand, B^{D} can be computed as follows:

At first, we consider

$$\begin{pmatrix} I & X \\ 0 & I \end{pmatrix} \begin{pmatrix} J & E_0 \\ 0 & N \end{pmatrix} \begin{pmatrix} I & -X \\ 0 & I \end{pmatrix} = \begin{pmatrix} J & 0 \\ 0 & N \end{pmatrix}.$$

where X is determined by the Sylvester equation [18] $JX - XN = E_0$. Let $X = [x_1, x_2, \ldots, x_k]$, $E_0 = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k]$. Then the above equation is equivalent to

$$(I \otimes J - N^{\mathrm{T}} \otimes I)\mathrm{vec}(X) = \mathrm{vec}(E_0),$$

i.e.,

$$\begin{pmatrix} J & & & \\ -I & J & & \\ & \ddots & \ddots & \\ & & -I & J \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{pmatrix} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_k \end{pmatrix} \Rightarrow \begin{cases} x_1 = J^{-1}\varepsilon_1, \\ x_2 = J^{-2}\varepsilon_1 + J^{-1}\varepsilon_2, \\ \vdots \\ x_k = J^{-k}\varepsilon_1 + \cdots + J^{-1}\varepsilon_k, \end{cases}$$

Thus

$$B = \begin{pmatrix} I & -X \\ 0 & I \end{pmatrix} \begin{pmatrix} J & 0 \\ 0 & N \end{pmatrix} \begin{pmatrix} I & X \\ 0 & I \end{pmatrix},$$

$$B^{\mathrm{D}} = \begin{pmatrix} I & -X \\ 0 & I \end{pmatrix} \begin{pmatrix} J^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} J^{-1} & J^{-1}X \\ 0 & 0 \end{pmatrix},$$

and so

$$B^{\mathrm{D}} - A^{\mathrm{D}} = \begin{pmatrix} 0 & J^{-1}X \\ 0 & 0 \end{pmatrix}.$$

This example illustrates the sensitivity of the Drazin inverse depends on the quantity $\|(A^{\mathrm{D}})^k\|$ generally.

Furthermore, we believe that the condition number in the definition [22] is not only complicated but also overestimate. Later, Wei and Wang [29] gave a simple bound for $\frac{\|B^{\mathrm{D}} - A^{\mathrm{D}}\|}{\|A^{\mathrm{D}}\|}$ under the condition $B = A + E$ and $E = AA^{\mathrm{D}}EAA^{\mathrm{D}}$. However, this is a kind of the special perturbation. In fact, the Drazin inverse can be represented explicitly by means of the Jordan canonical form for A as follows. Let

$$A = P \begin{pmatrix} C & 0 \\ 0 & N \end{pmatrix} P^{-1}, \tag{3.2}$$

where P is nonsingular matrix, C is a nonsingular upper bidiagonal matrix and N is nilpotent of index k, that is, $N^k = 0$ and $N^{k-1} \neq 0$. Then

$$A^{\mathrm{D}} = P \begin{pmatrix} C^{-1} & 0 \\ 0 & 0 \end{pmatrix} P^{-1}, \; AA^{\mathrm{D}} = P \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} P^{-1}. \tag{3.3}$$

Thus $E = AA^{\mathrm{D}}EAA^{\mathrm{D}}$ if and only if

$$E = P \begin{pmatrix} E_0 & 0 \\ 0 & 0 \end{pmatrix} P^{-1}. \tag{3.4}$$

In addition, the main results of [29] can be proved by using (3.3) and (3.4) directly.

To sum up, further researches about the condition number and perturbation analysis of Drazin inverse may be of some significance.

Recently, Li and Wei [15] gave a perturbation bound for the group inverse which can be obtained by using another expression of Drazin inverse and the results from Li and Wei [16].

We will need some lemmas.

Lemma 3.2. *Let* $A \in \mathbf{C}^{n\times n}$, *with* $\mathrm{Ind}(A) = k$. *Then* $(A^k)^\sharp$ *exists, and*

$$A^{\mathrm{D}} = A^{k-1}(A^k)^\sharp. \tag{3.5}$$

Moreover

$$AA^{\mathrm{D}} = A^k(A^k)^\sharp. \tag{3.6}$$

Using (3.2) and (3.3), we can obtain (3.5) and (3.6) directly. This is another expression of Drazin inverse A^{D}.

The following lemma gives a necessary condition for the continuity of the Drazin inverse.

Lemma 3.3. *Let* $A, B = A + E \in \mathbf{C}^{n\times n}$, $k = \max\{\mathrm{Ind}(B), \mathrm{Ind}(A)\}$. *If* $\mathrm{rank}(B^k) > \mathrm{rank}(A^k)$, *then*

$$\|B^{\mathrm{D}}\|_2 \geqslant \frac{1}{\|B^k - A^k\|_2\|(B^k)^\sharp B\|_2}, \tag{3.7}$$

and

$$\|B^{\mathrm{D}}\|_2^k \geqslant \frac{1}{\|B^k - A^k\|_2\|(B^k)^\sharp B^k\|_2}. \tag{3.8}$$

Proof. Since $\mathrm{rank}(B^k) > \mathrm{rank}(A^k)$, there exists a $y \in \mathbf{C}^n$ such that

$$\|y\|_2 = 1, \quad y^H A^k = 0, \quad y \in \mathscr{R}(B^k).$$

Notice that BB^{D} is the oblique projector along $\mathscr{N}(B)$ onto $\mathscr{R}(B)$, and $(B^k)^\sharp = B^{\mathrm{D}}(B^k)^\sharp B$. For

$$\begin{aligned} 1 &= yy^H = y^H BB^{\mathrm{D}}y = y^H B^k(B^k)^\sharp y \quad \text{(see Lemma 3.2)} \\ &= y^H(B^k - A^k)(B^k)^\sharp y = y^H(B^k - A^k)B^{\mathrm{D}}(B^k)^\sharp By \\ &\leqslant \|B^k - A^k\|_2\|(B^k)^\sharp B\|_2\|B^{\mathrm{D}}\|_2. \end{aligned}$$

Then the conclusion (3.7) follows. By $(B^k)^\sharp = (B^{\mathrm{D}})^k(B^k)^\sharp B^k$, the inequality (3.8) is obtained similarly. □

The two lemmas above make it clear that to obtain useful perturbation bounds we should assume $\mathrm{rank}(B) = \mathrm{rank}(A)$ for the group inverse.

The following result is relatively weaker than the one by Li and Wei ([15, Theorem 3]), but more explicit.

Lemma 3.4. *Let* $B = A + E$, $\mathrm{Ind}(A) \leqslant 1$. *If* $\mathrm{rank}(B) = \mathrm{rank}(A)$, *and*

$$\|A^\sharp\|_2\|E\|_2 < \frac{1}{1 + \mathrm{Ind}(A)\sqrt{\|AA^\sharp\|_2}},$$

then $\mathrm{Ind}(B) \leqslant 1$, *and*

$$\frac{\|B^\sharp - A^\sharp\|_2}{\|A^\sharp\|_2} \leqslant \frac{(1+v_1)(1+v_2)}{(1-\|A^\sharp\|_2\|E\|_2)(1-v_1v_2)^2} - 1,$$

where

$$v_1 = \frac{\|A^\sharp\|_2\|E(I-AA^\sharp)\|_2}{1-\|A^\sharp\|_2\|E\|_2}, \qquad v_2 = \frac{\|A^\sharp\|_2\|(I-AA^\sharp)E\|_2}{1-\|A^\sharp\|_2\|E\|_2}.$$

The main results of this section are as follows.

Theorem 3.5. *Let* $A, B \in \mathbf{C}^{n\times n}$, $k = \mathrm{Ind}(A)$, *if* $\mathrm{rank}(B^k) = \mathrm{rank}(A^k)$, $E_k \equiv B^k - A^k$ *and*

$$\|(A^k)^\sharp\|_2\|E_k\|_2 < \frac{1}{1+\sqrt{\|A^k(A^k)^\sharp\|_2}}. \tag{3.9}$$

Then $\mathrm{Ind}(B) \leqslant k$, *and*

$$\frac{\|B^{\mathrm{D}} - A^{\mathrm{D}}\|_2}{\|A^{k-1}\|_2\|(A^k)^\sharp\|_2} \leqslant \frac{\|B^{k-1}\|_2}{\|A^{k-1}\|_2}\left[\frac{(1+\varDelta_1)(1+\varDelta_2)}{(1-\|(A^k)^\sharp\|_2\|E_k\|_2)(1-\varDelta_1\varDelta_2)} - 1\right] + \frac{\|B^{k-1} - A^{k-1}\|_2}{\|A^{k-1}\|_2}, \tag{3.10}$$

where

$$\varDelta_1 = \frac{\|(A^k)^\sharp\|_2\|E_k(I - A^k(A^k)^\sharp)\|_2}{1-\|(A^k)^\sharp\|_2\|E_k\|_2}, \tag{3.11}$$

and

$$\varDelta_2 = \frac{\|(A^k)^\sharp\|_2\|(I - A^k(A^k)^\sharp)E_k\|_2}{1-\|(A^k)^\sharp\|_2\|E_k\|_2}. \tag{3.12}$$

Proof. From $\mathrm{Ind}(A) = k$, we have $\mathrm{Ind}(A^k) \leqslant 1$. From Lemma 3.4, $\mathrm{Ind}(B^k) \leqslant 1$. Using the Jordan canonical form for B, we know $\mathrm{Ind}(B) \leqslant k$, so $(B^k)^\sharp$ exists. By Lemma 3.2 we obtain

$$B^{\mathrm{D}} - A^{\mathrm{D}} = B^{k-1}(B^k)^\sharp - A^{k-1}(A^k)^\sharp = B^{k-1}[(B^k)^\sharp - (A^k)^\sharp] + (B^{k-1} - A^{k-1})(A^k)^\sharp.$$

Taking the 2-norm, we obtain (3.10). □

As $AA^{\mathrm{D}} = (A^k)(A^k)^\sharp$, We can derive a perturbation bound for the oblique projection AA^{D} by using a result of Li and Wei [15, Theorem 6].

Theorem 3.6. *Let* $A, B \in \mathbf{C}^{n\times n}$, $\mathrm{Ind}(A) = k$, $E_k = B^k - A^k$. *If* $\mathrm{rank}(B^k) = \mathrm{rank}(A^k)$ *and*

$$\|(A^k)^\sharp\|_2\|E_k\|_2 < \frac{1}{1+\sqrt{\|A^k(A^k)^\sharp\|_2}}, \tag{3.13}$$

then $\mathrm{Ind}(B) \leqslant k$, $\|y\|_2 < 1$, *and*

$$\|BB^{\mathrm{D}} - AA^{\mathrm{D}}\|_2 \leqslant \frac{\|(A^k)^\sharp\|_2\|AA^{\mathrm{D}}\|_2\|E_k\|_2}{(1-\|y\|_2)(1-\|(A^k)^\sharp\|_2\|E_k\|_2)^2} + \frac{\|(I-AA^{\mathrm{D}})E_k\|_2\|(A^k)^\sharp\|_2^2\|E_k(I-AA^{\mathrm{D}})\|_2}{(1-\|y\|_2)(1-\|(A^k)^\sharp\|_2\|E_k\|_2)^2}, \tag{3.14}$$

where

$$y = (A^k)^\sharp (I + E_k (A^k)^\sharp)^{-1} E_k (I - A^k (A^k)^\sharp)(I + (A^k)^\sharp E_k)^{-1} (A^k)^\sharp. \tag{3.15}$$

Noting that, if $\|(A^k)^\sharp\|_2 \|E_k\|_2$ and $\varDelta_1, \varDelta_2$ are small, then in the first order approximation, we have

$$\frac{(1+\varDelta_1)(1+\varDelta_2)}{(1-\|(A^k)^\sharp\|_2\|E_k\|_2)(1-\varDelta_1\varDelta_2)} = 1 + \varDelta_1 + \varDelta_2 + \|(A^k)^\sharp\|_2 \|E_k\|_2,$$

where

$$\varDelta_1 = \|A^k\|_2 \|(A^k)^\sharp\|_2 \frac{\|E_k(I - A^k(A^k)^\sharp)\|_2}{\|A^k\|_2} \frac{1}{1 - \|A^k\|_2 \|(A^k)^\sharp\|_2 \frac{\|E_k\|_2}{\|A^k\|_2}},$$

$$\varDelta_2 = \|A^k\|_2 \|(A^k)^\sharp\|_2 \frac{\|(I - A^k(A^k)^\sharp)E_k\|_2}{\|A^k\|_2} \frac{1}{1 - \|A^k\|_2 \|(A^k)^\sharp\|_2 \frac{\|E_k\|_2}{\|A^k\|_2}},$$

and

$$\|(A^k)^\sharp\|_2 \|E_k\|_2 = \|A^k\|_2 \|(A^k)^\sharp\|_2 \frac{\|E_k\|_2}{\|A^k\|_2}.$$

In general, (3.10) and (3.14) cannot be regarded as an improvement of the known results, but it complements them. Thus the condition number can be given as follows

$$\kappa_d \equiv \|A^k\|_2 \|(A^k)^\sharp\|_2. \tag{3.16}$$

Notice that, Rong [22], Wei and Wang [29], Wei and Wu [31], Campbell and Mayer [3] have defined the condition numbers in different situations. In fact, each of the condition numbers (including κ_d) can be used in appropriate cases respectively.

Example 3.5. Let

$$A = \operatorname{diag}(I_2, J_3(0)), \quad J_3(0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & \varepsilon & 0 \\ 0 & 0 & \varepsilon \\ 0 & 0 & 0 \end{pmatrix},$$

$$B = \operatorname{diag}(I_2, J_3(0) + E_3).$$

It is easy to see that $\operatorname{Ind}(A) = \operatorname{Ind}(B) = 2$, $\kappa_d = 1$, $\varDelta_1 = \varDelta_2 = 0$, $\|B^{\mathrm{D}} - A^{\mathrm{D}}\|_2 = 0$, but by [27, p.228, (8.4.1)] we have estimate

$$C(A) = 9, \ \frac{\|B^{\mathrm{D}} - A^{\mathrm{D}}\|_2}{\|A^{\mathrm{D}}\|_2} \lesssim 9|\varepsilon|.$$

Following the approach by Li and Wei [15], the condition number κ_d can be estimated as follows.

Let the Schur decomposition of A [13] be given by

$$A = Q \begin{pmatrix} B & D \\ 0 & C \end{pmatrix} Q^H,$$

where Q is a unitary matrix, B is nonsingular, and the diagonal elements of C are zeros. Consider the following matrix

$$X_0 = \begin{pmatrix} I & K \\ 0 & I \end{pmatrix}, \quad BK - KC = D.$$

Let $Q_0 = Q\begin{pmatrix} I & -K \\ 0 & I \end{pmatrix}$. Then we have

$$A = Q_0 \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix} Q_0^{-1}.$$

It follows from $\mathrm{Ind}(A) = k$ that

$$A^k = Q_0 \begin{pmatrix} B^k & 0 \\ 0 & 0 \end{pmatrix} Q_0^{-1}, \quad (A^k)^\sharp = Q_0 \begin{pmatrix} (B^k)^\sharp & 0 \\ 0 & 0 \end{pmatrix} Q_0^{-1} = Q_0 \begin{pmatrix} B^{-k} & 0 \\ 0 & 0 \end{pmatrix} Q_0^{-1}.$$

Thus

$$\kappa_d \leqslant (\|Q_0\|_2 \|Q_0^{-1}\|_2)^2 \|B^k\|_2 \|B^{-k}\|_2,$$

and

$$\|Q_0\|_2^2 = 1 + \frac{1}{2}\left[\sigma_1^2 + \sigma_1\sqrt{4+\sigma_1^2}\right], \quad \|Q_0^{-1}\|_2^2 = 1 + \frac{1}{2}\left[\sigma_2^2 - \sigma_2\sqrt{4+\sigma_2^2}\right],$$

where σ_1 and σ_2 are the biggest and smallest singular values of K, respectively.

Acknowledgements

The authors thank Prof. H. Bart, R. Brualdi and the referee for their very useful and detailed comments which greatly improve the presentation.

References

[1] A. Ben-Israel, T.N.E. Greville, Generalized Inverse: Theory and Applications, John Wiley, New York, 1974; A. Ben-Israel, T.N.E. Greville, Generalized Inverse: Theory and Applications, second ed., Springer, 2003.
[2] S.L. Campbell, C.D. Meyer, Generalized Inverses of Linear Transformations, Pitman, London, 1979.
[3] S.L. Campbell, C.D. Meyer, Continuity properties of the Drazin inverse, Linear Algebra Appl. 10 (1975) 77–83.
[4] S.L. Campbell, On continuity properties of the Moore–Penrose and Drazin generalized inverse, Linear Algebra Appl. 18 (1977) 53–57.
[5] N. Castro González, J.J. Koliha, Y. Wei, Perturbation of Drazin inverse for matrices with equal eigenprojections at zero, Linear Algebra Appl. 312 (2000) 181–189.
[6] N. Castro González, J.J. Koliha, I. Straškraba, Perturbation of the Drazin inverse, Soochow J. Math. 27 (2001) 201–211.
[7] N. Castro González, J.J. Koliha, V. Rakočević, Continuity and general perturbation of the Drazin inverse for closed linear operators, Abstract Appl. Anal. 7 (2002) 335–347.
[8] N. Castro González, J.J. Koliha, New additive results for the g-Drazin inverse, Proc. Roy. Soc. Edinburgh, Sect. A 134 (2004) 1085–1097.
[9] N. Castro González, J. Vlez-Cerrada, Characterizations of matrices which eigenprojections at zero are equal to a fixed perturbation, Appl. Math. Comput. 159 (2004) 613–623.
[10] N. Castro González, Additive perturbation results for the Drazin inverse, Linear Algebra Appl. 397 (2005) 279–297.
[11] D. Djordjević, Y. Wei, Additive results for the generalized Drazin inverse, J. Aust. Math. Soc. 73 (2002) 115–125.
[12] E. Eiermann, I. Marek, W. Niethammer, On the solution of singular linear equations by semiiterative methods, Numer. Math. 53 (1988) 265–283.
[13] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., John Hopkins University Press, Baltimore, MD, 1996.
[14] R. Hartwig, G. Wang, Y. Wei, Some additive results on Drazin inverse, Linear Algebra Appl. 322 (2001) 207–217.
[15] X. Li, Y. Wei, An improvement on the perturbation of the group inverse and oblique projection, Linear Algebra Appl. 338 (2001) 53–66.
[16] X. Li, Y. Wei, A note on the perturbation bound of the Drazin inverse, Appl. Math. Comput. 140 (2003) 329–340.
[17] J.J. Koliha, Error bounds for a general perturbation of the Drazin inverse, Appl. Math. Comput. 126 (2002) 181–185.
[18] M. Konstaintinov, D. Gu, V. Mehrmann, P. Petkov, Perturbation Theory for Matrix Equations, Elsevier, 2003.

[19] C.D. Meyer, The role of the group generalized inverse in the theory of finite Markov chain, SIAM Rev. 17 (1975) 443–464.

[20] V. Rakočević, Y. Wei, The perturbation theory for the Drazin inverse and its applications II, J. Aust. Math. Soc. 70 (2001) 189–197.

[21] V. Rakočević, Koliha-Drazin inverse and perturbations, Rend. Circ. Mat. Palermo (2, Suppl. 73) (2004) 101–125.

[22] G.H. Rong, The error bound of the perturbation of the Drazin inverse, Linear Algebra Appl. 47 (1982) 159–168.

[23] J.-G. Sun, Matrix Perturbation Analysis, second ed., Science Press, Beijing, 2001 (in Chinese).

[24] J.-G. Sun, Backward perturbation analysis of certain characteristic subspaces, Numer. Math. 65 (1993) 357–382.

[25] G.W. Stewart, On the continuity of the generalized inverse, SIAM J. Appl. Math. 17 (1969) 33–45.

[26] G.W. Stewart, On the of the perturbation of pseudo-inverse, projections and linear least squares problems, SIAM Rev. 19 (1977) 634–662.

[27] G. Wang, Y. Wei, S. Qiao, Generalized Inverses: Theory and Computations, Science Press, Beijing, 2004.

[28] P.Å. Wedin, Perturbation theory for pseudoinverses, BIT 13 (1973) 217–232.

[29] Y. Wei, G. Wang, The perturbation theory for the Drazin inverse and its applications, Linear Algebra Appl. 258 (1997) 179–186.

[30] Y. Wei, On the perturbation of the group inverse and oblique projection, Appl. Math. Comput. 98 (1999) 29–42.

[31] Y. Wei, H. Wu, The perturbation of the Drazin inverse and oblique projection, Appl. Math. Lett. 13 (2000) 77–83.

[32] Y. Wei, H. Wu, Challenging problems on the perturbation of Drazin inverse, Ann. Oper. Res. 103 (2001) 371–378.

[33] Y. Wei, Perturbation bound of the Drazin inverse, Appl. Math. Comput. 125 (2002) 231–244.

[34] Y. Wei, X. Li, An improvement on perturbation bounds for the Drazin inverse, Numer. Linear Algebra Appl. 10 (2003) 563–575.

[35] Y. Wei, X. Li, F. Bu, A perturbation bound of the Drazin inverse of a matrix by separation of simple invariant subspaces, SIAM J. Matrix Anal. Appl. 27 (2005) 72–81.

[36] Y. Wei, X. Li, F. Bu, F. Zhang, Relative perturbation bounds for the eigenvalues of diagonalizable and singular matrices – Application of perturbation theory for simple invariant subspaces, Linear Algebra Appl. 419 (2006) 765–771.